# EXHIBIT D

TO DECLARATION OF S. MERRILL WEISS IN SUPPORT OF PLAINTIFF ACACIA MEDIA TECHNOLOGIES CORPORATION'S MEMORANDUM OF POINTS AND AUTHORITIES IN OPPOSITION TO ROUND 3 DEFENDANTS' MOTION FOR SUMMARY JUDGMENT OF INVALIDITY UNDER 35 U.S.C. § 112 OF THE '992, '863, AND '702 PATENTS; AND SATELLITE DEFENDANTS' MOTION FOR SUMMARY JUDGMENT OF INVALIDITY OF THE '992, '863, AND '720 PATENTS

Hovsepian v. Apple, Inc.

Doc. 325 Att. 3

Dockets.Justia.com

RCA 86,314

AR CITED BY APPLICANT

# A CHIP SET CORE FOR IMAGE COMPRESSION

Alain Artieri, Oswald Colavin

SGS-THOMSON Microelectronics, Image Processing Business Unit
17, avenue des Martyrs - B.P. 217
38019 GRENOBLE Cedex FRANCE

## Abstract

Todays efficient codecs for full motion image compression are based on two key functions : Discrete Cosine Transform (DCT) and Motion Estimation. This fact is strongly supported by international standardization committees (ISO,CCITT) that are currently standardizing techniques for compression of pictures (still and moving) based on DCT and Motion Compensation.
Two new components performing these functions are presented. The first component computes 8*8 Discrete Cosine Transform and zig zag conversion of coefficient scanning for a pixel rate up to 27 Mhz. The second component computes full search Motion Estimation for a pixel rate up to 18 Mhz.
System implementation for image compression is then discussed.

## Introduction

Image compression techniques are of increasing economical interest as image communication comes on the market. A wide range of applications are directly concerned by compression :

- Image Communication on ISDN (H261 standard).
- Digital Storage Media (DSM) including CD Rom, digital VCR, etc...
- TV Transmission (Contribution, Distribution).
- Cable TV
- HDTV

The techniques that exhibit the best tradeoff between picture quality and system cost are based on two fundamental functions : Discrete Cosine Transform (DCT) and Motion Compensation.
However, real time Compression and Decompression of motion video requires a computation power which is not compatible with classical general purpose DSP. For example, compression of CCIR 601 picture requires about 7,000,000,000 operations per second !!! Only dedicated VLSI can deal with such computation power.
Two new VLSI circuits for computing DCT and Motion estimation are presented in this communication followed by a short discussion on system aspects.

## Algorithm overview

A compression algorithm takes advantage of the spatial redundancy in a picture and of the redundancy between successive pictures. Instead of processing single pixels, efficient algorithms consider blocks of pixel.
The first operation consists in segmenting each picture into blocks. Each block may be encoded by an INTRA technique or a PREDICTIVE technique.

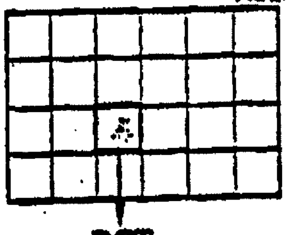

figure 1 : INTRA coding

The INTRA coding consists in applying first a two dimensional transform to the block. The result is a block of numbers called coefficients. The human eye has not the same sensitivity with coefficients that are at different positions in the block. As a result, each coefficient is quantized with a quantized step related to the eye sensitivity. After these operations, a block

AMTC015312

of pixels is reduced to a very few number of non zero coefficient that are then encoded with an entropic coder (Huffman coder, Q coder).

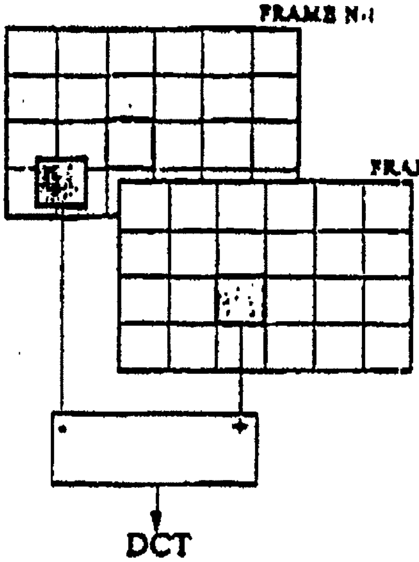

figure 2 : PREDICTIVE coding

The PREDICTIVE coding consists in encoding the difference between the block and a prediction of this block. The prediction of the block is a displaced block in the previous picture. The displacement is computed by a Motion Estimator and is generally related to the motion in the scene. The maximum displacement is limited in order to restrain the search in a local area of the previous picture around the block. The error block is then transformed with a DCT and quantized. The non zero coefficient and the displacement are encoded with the entropic coder.

A third technique has been proposed that leads to very good results : the INTERPOLATIVE coding. This technique is similar to the predictive coding, but two displaced blocks are found : one in the previous frame and one in the next frame. The best predictor is then chosen between the displaced block in the previous frame, the displaced block in the next frame and the average of the two.

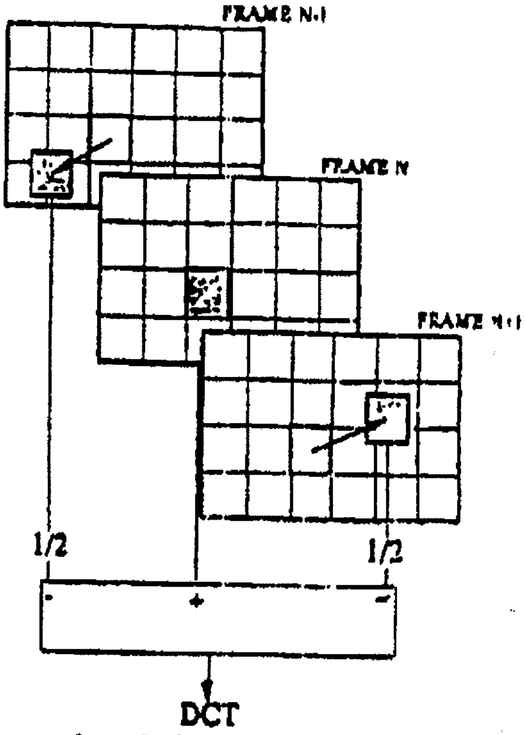

figure 3 : INTERPOLATIVE coding

A combination of the three techniques allows compression at 1.2 Mbit/sec with a picture quality corresponding to VHS quality as shown by recent results of the MPEG standardization committee.
For a 4:1:1 sampling scheme, the computation power of the different steps are :

| | |
|---|---|
| DCT : | 15 Op/pixel |
| QUANTISATION : | 1 Op/pixel |
| MOTION ESTIMATION : | 512 Op/pixel (+7/-8)<br>2,048 Op/pixel (+15/-16) |

As a result, the engine of a compression system is made of the two functions : DCT and Motion estimation.

AMTC015313

Exhibit D Page 157

## Discrete Cosine Transform

DCT is a two dimensional operation that is performed on blocks of data. The most commonly used block format is 8*8 pixels. In the transformed space, data are scanned in a zig zag order (figure 4) in order to reduce the entropy. For these reasons, we have developed a device that computes 8*8 DCT and includes the zig zag scanning conversion of coefficient block.

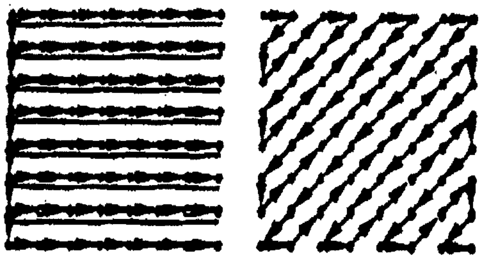

figure 4 : ZIG ZAG scanning of DCT coefficients

The accuracy of the computation is an important feature, specially for the inverse DCT when a coding loop is used. As a part of the H261 standard, the CCITT WG15 has proposed a standard for inverse DCT accuracy and we have developed a new algorithm for this purpose.
This new algorithm (figure 5) uses 11 multiplications and 29 additions to compute a one dimensional 8 point DCT : this is the minimum known nowadays.

The three main building blocks are the operative part, the transposition memory and the zig zag scanning conversion block (figure 6).

Operative part :

The operative part is a direct mapping of the new fast flow graph onto silicon. The cosine constants are coded with 12 bits accuracy.

Transposition memory :

A dedicated register bank has been preferred to a classical RAM structure as a low memory size is required (1,024 bits).

Zig Zag scanning conversion :

The zig zag scanning conversion block is also a dedicated register bank with the result that the latency is only 30 cycles. A classical RAM based architecture would have led to 64 cycles.
This block is a set of 27 modules. Each module (figure 7) performs a simple permutation of each set of 64 coefficients. The combination of the 27 permutations forms the Zig Zag permutation.
Each module includes a register and a selector. A set of 64 predefined commands define the permutation for the module.
Three pipe line registers have been added in order to reduce the critical path resulting in a total latency of 30 cycles.

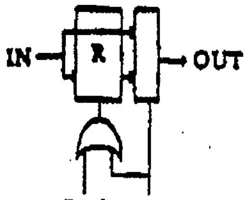

figure 7 : ZigZag module

Forward/Inverse Transform management

The architecture is fully bidirectional. Each building block performs the forward function when the data flow in one direction and the inverse function when data flow in the reverse direction.

Precision Issue

We propose two precision modes. The single precision mode allows high data rate processing (up to 27 Mhz). With this mode, the internal data format of the operative part is 16 bits.The double precision mode improves the accuracy of computation with a lower maximum data rate (up to 20 Mhz). This mode is handled by doubling the operative part clock frequency with the result that the internal data format of the operative part becomes 32 bits.

AMTC015314

Exhibit D Page 158

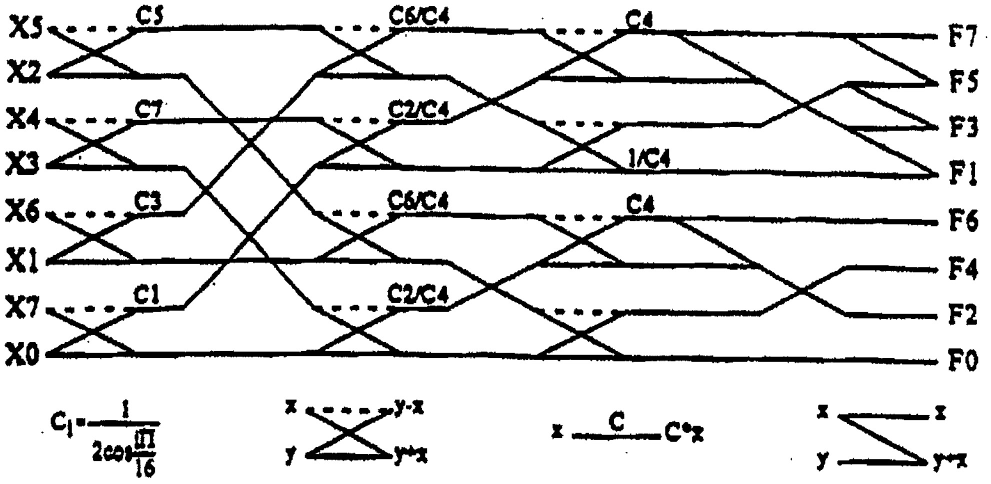
X5
X2
X4
X3
X6
X1
X7
X0
C5
C7
C3
C1
C6/C4
C2/C4
C6/C4
C2/C4
C4
1/C4
C4
F7
F5
F3
F1
F6
F4
F2
F0
x
y
y-x
y+x
x
C
C*x
x
y
x
y+x

figure 5 : new fast flow graph

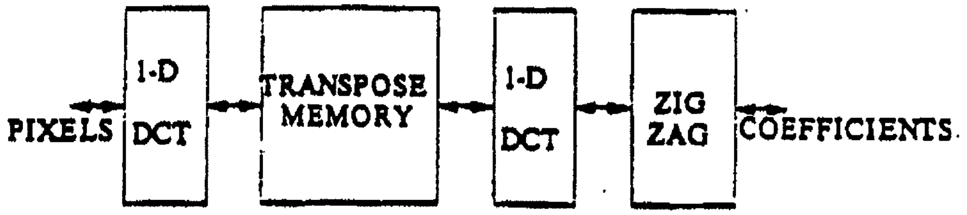
PIXELS
1-D
DCT
TRANSPOSE
MEMORY
1-D
DCT
ZIG
ZAG
COEFFICIENTS

figure 6 : DCT chip architecture

Chip characteristics :

| | |
|---|---|
| number of transistors : | 86,008 |
| chip size : | 30mm2 |
| technology : | CMOS 1.2 micron |
| power dissipation : | 500 mW max |

The chip has been fabricated and tested and is fully functional.

## Motion Estimation

We have developed an architecture for the full search block matching algorithm which exhibits better results than non exhaustive algorithms.
For each block of a picture, a vector is computed that points to the best matching block in the previous picture. The search area is limited to a part of the picture around the position of the block.
The device tests 256 different positions corresponding to the vectors in the range +7/-8 pixels horizontally and vertically. A distortion is computed for each position according to the mean absolute error criterion and the minimum distortion and corresponding vector are outputted.

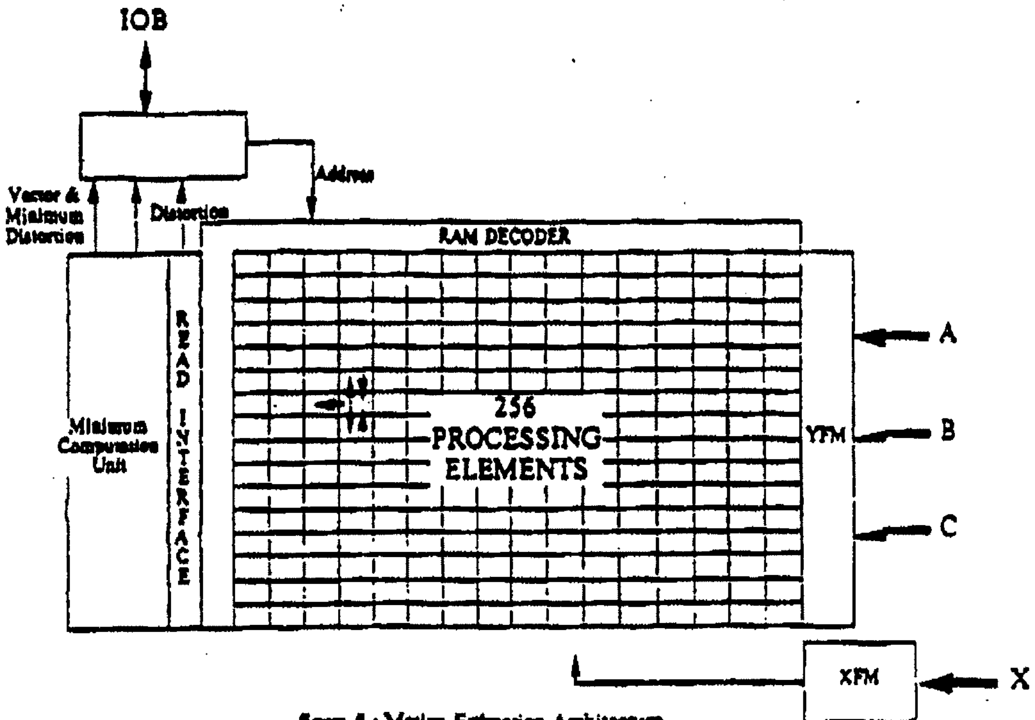
IOB
Vector & Minimum Distortion
Distortion
Address
RAM DECODER
Minimum Computation Unit
READ INTERFACE
256 PROCESSING ELEMENTS
YFM
A
B
C
XFM
X

figure 8 : Motion Estimation Architecture

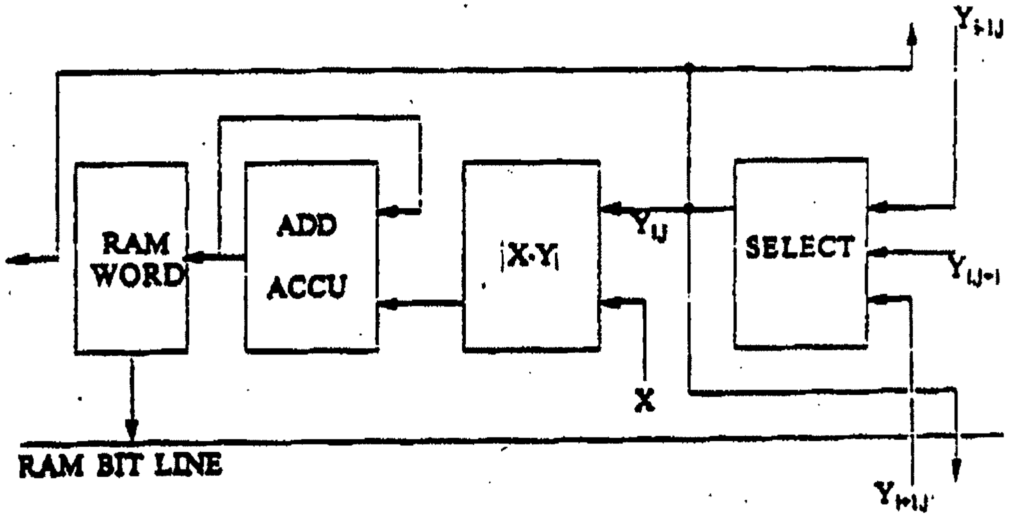
RAM WORD
ADD ACCU
|X-Y|
SELECT
X
RAM BIT LINE

figure 9 : Processing Element

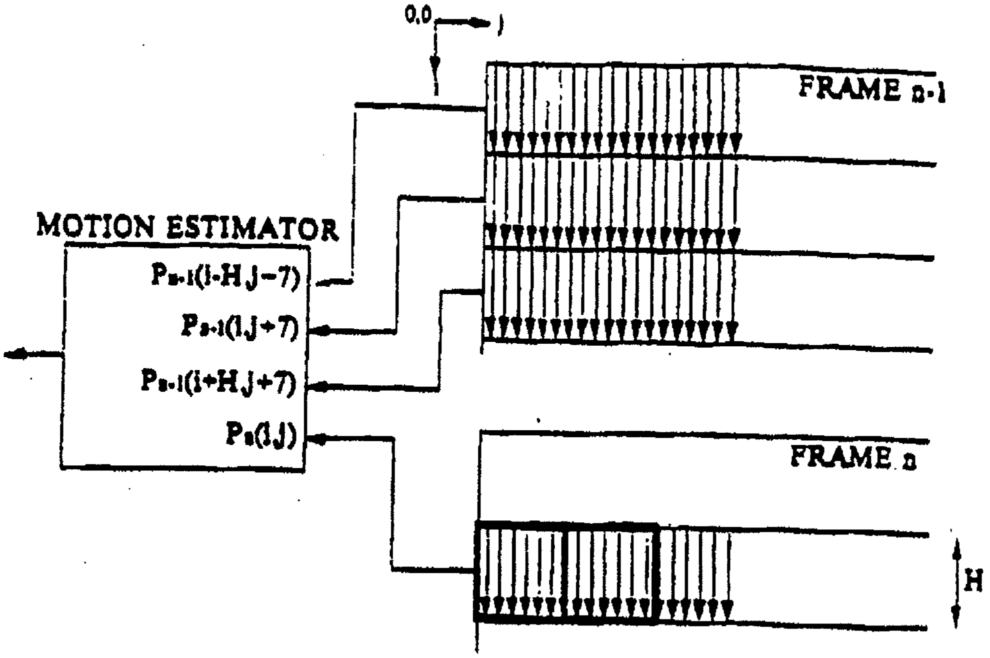

figure 10 : data scanning

A scalable pipe line architecture :

The architecture is based on a 16*16 systolic array (figure 8) of Processing Element (PE) (figure 9) that concurrently compute the 256 distortions. The architecture can process any block size but we have limited the possibilities to the most commonly used block size. The chip supports block heights of 8 and 16 pixels, and any block width by step of 4 pixels. For example, valid block sizes are 8*8, 8*16, 8*32, 16*8, 16*16...
The architecture is fully pipe lined with the result that the processing of a new block may start immediately after the last pixel a the preceding block is entered into the chip.
For example a pipe line processing of 8*8 blocks cost exactly 64 cycles, that is one cycle per pixel.

Random access to the distortions :

The 256 distortions values are stored in an on chip Random Access Memory mapped on the systolic array.

As a result, the distortions may be accessed by an external device. This allows to implement powerful motion compensation algorithms like sub pixel motion vector interpolation or motion vector field smoothing.

Data scanning :

The data must be scanned by stripe with a height equal to the block height. In order to meet the real time constraint, 4 pixels must be entered at each cycle. One pixel is for the current block and three pixels are for the search window (figure 10).

The maximum pixel rate is 18 Mhz. This allows to process picture format up to TV and HDTV format. Larger displacement computation may be handled with many chips (spatial multiplex) for higher pixel rates or with a single chip (time multiplex) for lower pixel rates.

AMTC015317

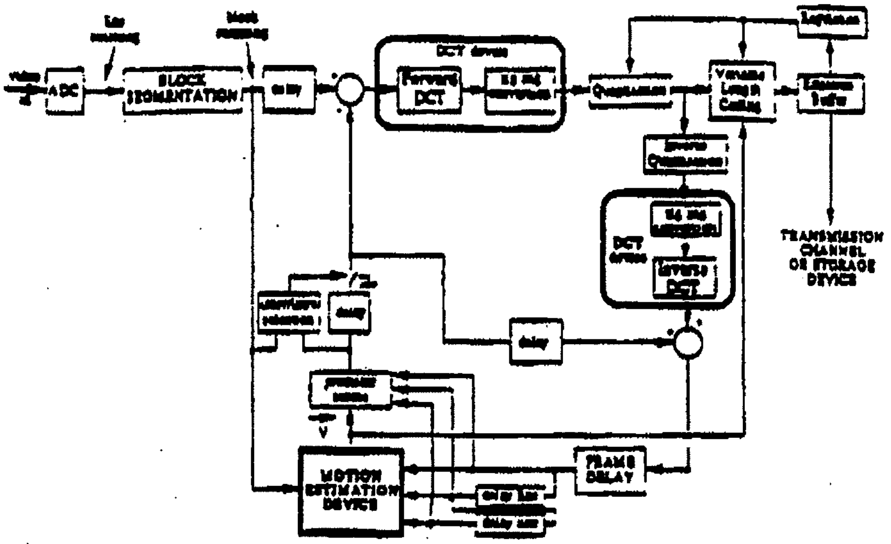
ADC
BLOCK SEGMENTATION
Forward DCT
DCT
FRAME DELAY
MOTION ESTIMATION DEVICE
TRANSMISSION CHANNEL OR STORAGE DEVICE

figure 11 : PIPE LINE ARCHITECTURE

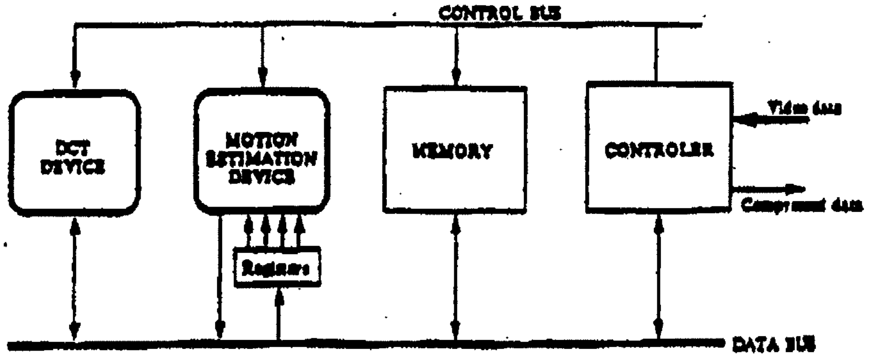
CONTROL BUS
DCT DEVICE
MOTION ESTIMATION DEVICE
Registers
MEMORY
CONTROLER
Video data
Compressed data
DATA BUS

figure 12 : TIME SHARING ARCHITECTURE

Chip characteristics :

| | |
|---|---|
| number of transistors : | 260,000 |
| chip size : | 70mm2 |
| technology : | CMOS 1.2 micron |
| power dissipation : | 2 W max |

The chip is currently designed and should be tested near October 1990.

## System Implementation

There are two main classes of codec architecture depending on the pixel rate.

Pipe line architectures :

For high pixel rates (TV,HDTV picture formats) a pipe line architecture will be preferred. For such rates, the bandwidth of each component is fully used and no time sharing between resources is possible. Figure 11 presents a conceptual pipe line architecture. The devices are cascaded and each one is synchronized by its predecessor and transmits a synchronization signal to its successor. The two proposed components include such synchronization features.

Time sharing architectures :

For lower pixel rates, resources may be shared in order to lower the system cost. A single memory for all the exchanges, a single DCT to compute all the forward and inverse DCT, a single motion estimator to compute larger displacements may be used and will lower the system cost. Figure 12 presents the corresponding conceptual architecture.

## Conclusion

Two components for image compression have been presented. They support almost all the huge computational power required by todays efficient compression algorithm. With these two components, cost effective solution for image compression may be developed.

## Acknowledgements

The authors would like to take this opportunity to acknowledge Francis Jutand and Nicolas Demassieux from TELECOM PARIS UNIVERSITY for their efficient contribution to the new architecture development and circuit design, and Yves Marie Le Pannerer from L.E.R. (Thomson Lab) for his helpful advises about the chips specification.

## References

H.G. Mussmann, P. Pirsch, and H.J. Grallert, "Advances in picture coding," Proc. IEEE, vol. 73, pp. 523-548, Apr. 1985.

A. Artieri, F. Jutand, "A Versatile and Powerfull Chip for Real Time Motion Estimation", ICASSP (Glasgow, 1989)

IEEE Transactions on Circuits and Systems, Vol. 36, No. 10, Oct. 1989.,